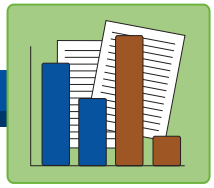
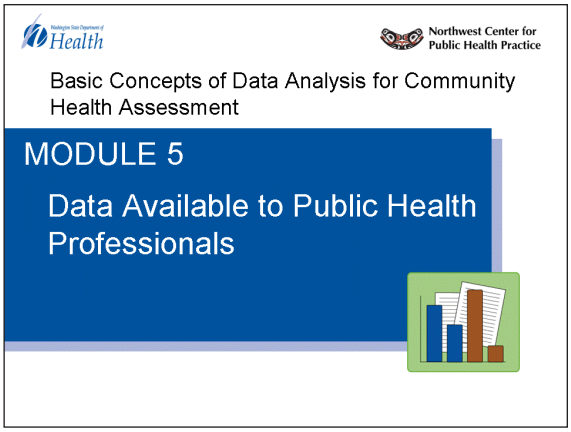
Module 5: Data Available to Public Health Professionals

Transcript

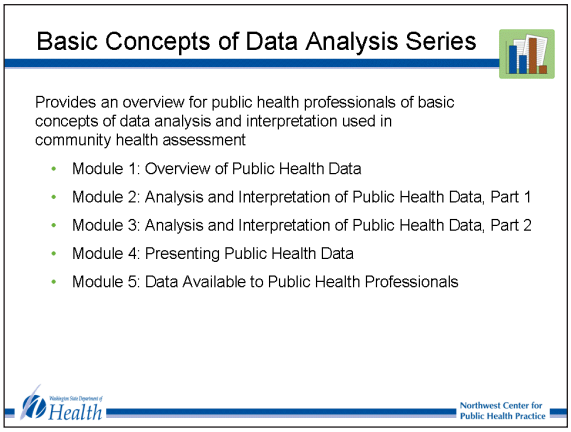


Data Available to Public Health Professionals in Washington State



Welcome to Data Available to Public Health Professionals in Washington State, the fifth module in the series on Basic Concepts of Data Analysis for Community Health Assessment in Washington State. I'm Jane Ballard. I helped create this course, along with a team of other community health assessment experts at both local and state health departments in Washington and Oregon. I have a doctorate in epidemiology and manage the Health Statistics and Assessment Program at Snohomish Health District. I have worked at Snohomish Health District since 2000.

Basic Concepts of Data Analysis Series



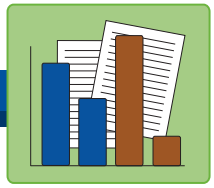
This series provides an overview for public health professionals of the basic concepts of data analysis and interpretation used in community health assessment. The training is intended to help professionals who work in public health practice at state and local agencies hone their assessment skills.

Module 1 provides an overview of public health data sources and uses. Module 2 introduces the analysis and interpretation of public health data. Module 3 continues the discussion of the analysis and interpretation of the public health data. Module 4 provides information on how to present public health data, and module 5 describes data available to public health professionals.

This series was developed by Washington State Department of Health in partnership with the Northwest Center for Public Health Practice. This module references both national and Washington State-specific sources of data. Other states have comparable offices that supply similar data and resources to those presented in this module. These offices may have different names, but can generally be found on the state's health department or other state government Web site.

Module 5: Data Available to Public Health Professionals

Transcript

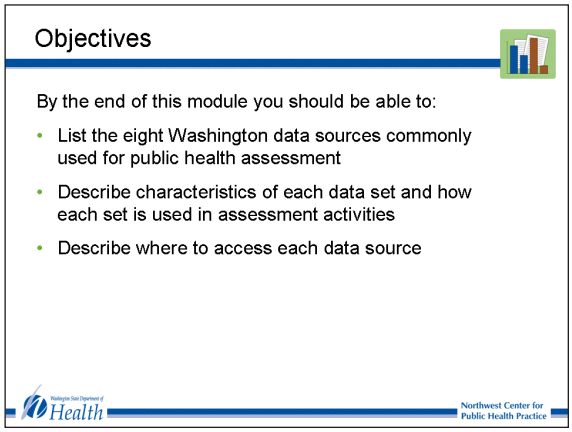


Objectives

In this module, we'll discuss data sources available to public health professionals in Washington State.

By the end of this module, you should be able to:

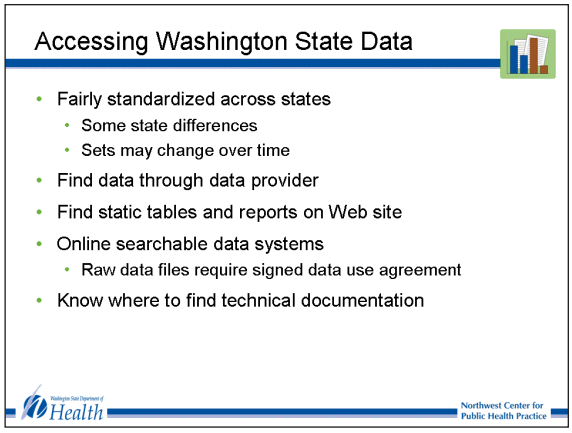
- List the eight Washington data sources commonly used for public health assessment
- Describe the characteristics of each data set and how each set is used in assessment activities, and
- Describe where to access each data source.



Accessing Washington State Data

The data sets we'll be looking at in this module are fairly standardized across the states. However, each state has some unique data collection, reporting practices, and processes for how data are shared. And, as with most data sets, these sets may change over time.

Although specific data access tools change over time, knowing the data provider for any particular data set will ensure that you can track down the information you need. Most data providers post static tables of data and reports on their Web sites that are updated annually. Some also have online searchable systems or will provide raw data files if you provide a signed data use agreement.



Finally, it's important to know where to find the technical documentation for a data set so you can clarify any questions that may come up as you work with the data set.

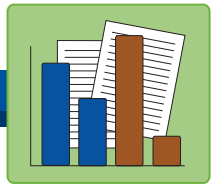
As we look at each data set, I'll discuss its collection methodology, data quality, variable definitions, confidentiality issues, geographic detail, and limitations.

Data Sources for Public Health Professionals

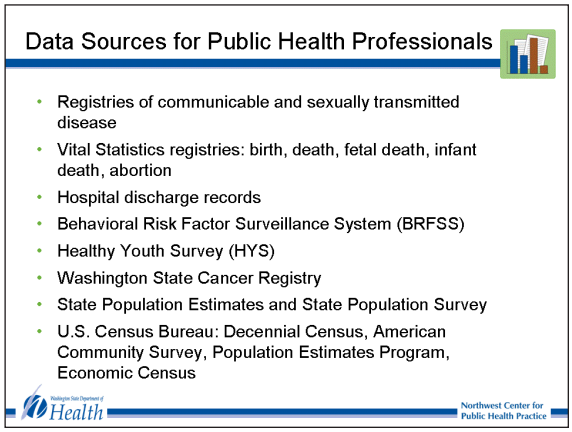
In this module we will discuss eight important data sets with information about Washingtonians:

- Registries of communicable and sexually transmitted diseases
- Vital statistic registries such as birth, death, fetal death, infant death, and abortion

Transcript



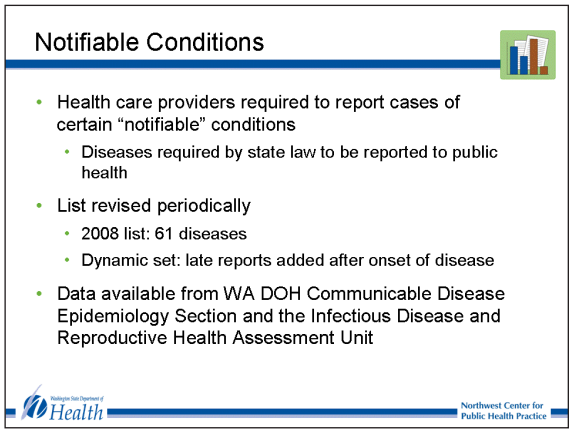
- Hospital discharge records
- Behavioral Risk Factor Surveillance System (or BRFSS)
- Healthy Youth Survey (or HYS)
- Washington State Cancer Registry
- Population Estimates and State Population Survey, and
- The U.S. Census Bureau's Decennial Census, American Community Survey, Population Estimates Program, and Economic Census



By the way, on the resources slide at the end of this module and in the attachments at the top of the screen, you'll find a list of these resources with information about where to access them.

Notifiable Conditions

Health care providers are required to report cases of certain "notifiable" communicable and sexually transmitted diseases. Notifiable conditions are diseases that are considered potentially dangerous to humans and must be reported to public health authorities at the time they're diagnosed. The list of nationally notifiable conditions is revised periodically. The 2008 list includes 61 diseases, such as measles, pertussis, *E. coli*, salmonella, tuberculosis, and various sexually transmitted diseases.



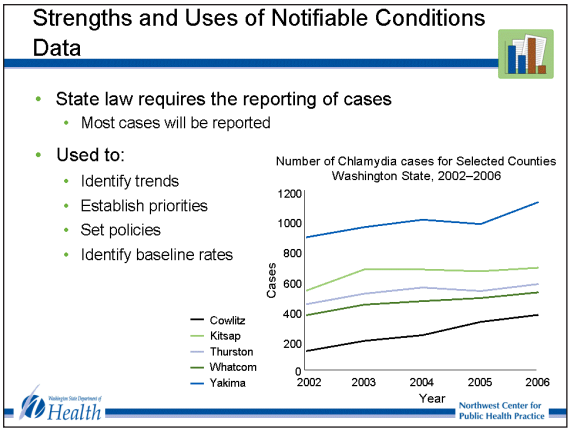
The communicable disease and sexually transmitted disease data are a dynamic data set with late reports being added for some time after the actual onset of the disease.

The Washington State Department of Health is responsible for statewide surveillance of these diseases under the Communicable Disease Epidemiology Section of the Office of Epidemiology and the Infectious Disease and Reproductive Health Assessment Unit in the Office of Infectious Disease and Reproductive Health.

Transcript

Strengths and Uses of Notifiable Conditions Data

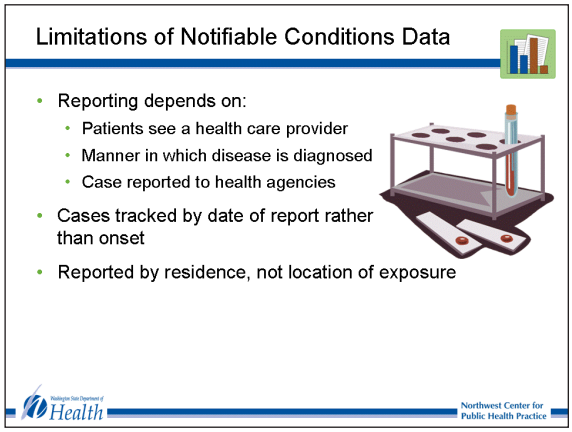
One of the strengths of communicable disease data is that state law requires the reporting of cases, increasing the likelihood that most cases will be reported. Both local and state health departments rely on the data to help them identify trends, establish priorities, and set policies. For example, because of the increasing trends in chlamydia in some counties, chlamydia interventions became a priority. The data are also used during outbreak investigations to identify baseline rates so the increase in cases can be quantified.



Limitations of Notifiable Conditions Data

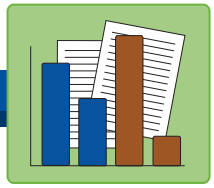
The communicable disease data have several limitations that are due the characteristics of diseases.

Reporting of "notifiable" diseases depends on three factors: the number of patients with severe enough symptoms that they see a health care provider, the manner in which the disease is diagnosed (for example, diagnosed by clinical or laboratory methods), and the case being reported to local health agencies. These factors vary according to the specific disease. As a result, data are more complete for some conditions, such as measles, than for others, such as salmonella.



Communicable disease reporting is most frequent with diseases that are severe or more likely to occur in children (since parents are more apt to seek medical care for even mild symptoms), because individuals with symptoms are more likely to be brought to the attention of a health care provider. Therefore, diseases such as meningococcal disease, measles, and *E. coli* in a child are much more likely to be reported than a slight case of salmonella in an adult. For the cause or diagnosis of these diseases to be confirmed, laboratory testing is necessary and labs generally meet the requirements to report to state authorities. However, clinicians still need to report the disease because labs do not provide important clinical data and labs are often not as timely in reporting cases.

Transcript



One other thing that is important to be aware of is that some local health departments track the number of cases by date of report rather than date of onset, which can result in differences in disease counts.

Finally, these data are reported by the residence of the patient, so they don't necessarily indicate where a person was exposed.

Let's pause now so you can check your understanding of the material we've just covered.

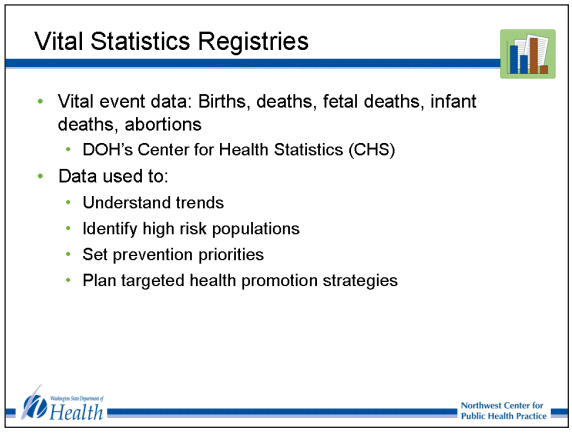
Practice: Notifiable Conditions Registries

Vital Statistics Registries

Vital event data, such as births, deaths, fetal deaths, infant deaths, and abortions, are usually maintained by the Department of Health's Center for Health Statistics.

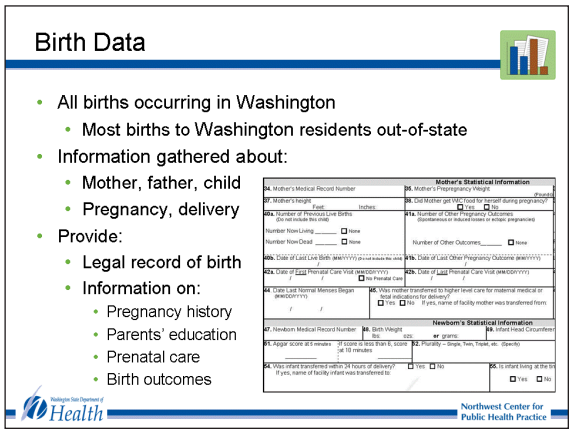
These data are widely used by policy makers, health professionals, community-based organizations, and health researchers to understand trends, identify high risk populations (and geographic areas), set prevention priorities, and plan targeted health promotion strategies.

Let's take a look at each of these sub-sets of vital records data.



Birth Data

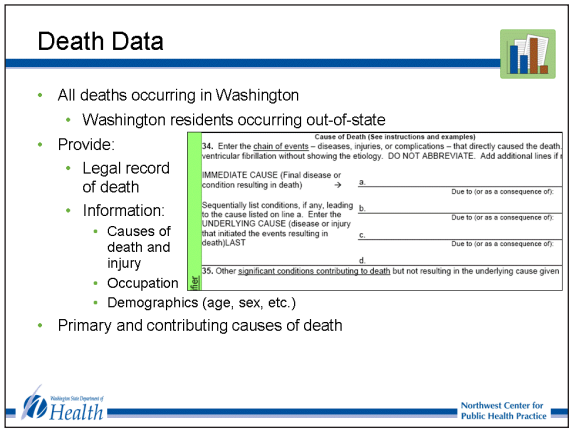
Birth data report all births occurring in Washington State, including residents and non-residents as well as most births to Washington State residents occurring out-of-state. Birth certificates gather information about the mother, father, child, and events of the pregnancy and delivery. They provide a legal record of the birth and valuable information for health assessment, such as pregnancy history, parents' education, prenatal care, mother's smoking during pregnancy and birth outcomes such as birth weight and gestation.



Transcript

Death Data

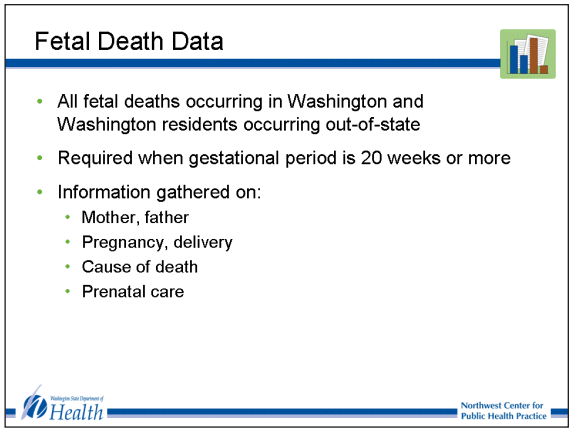
Death data report all deaths occurring in Washington State of both residents and non-residents as well as most deaths of Washington State residents occurring out-of-state. Death certificates provide a legal record of death and record information important for health assessment, such as causes of death and injury, occupation, age, and gender. Immediate, underlying and contributory causes of death may be reported on a death certificate. For example, although the immediate cause of death might be pneumonia, the underlying cause-of-death might be cancer or HIV or heart disease and contributing causes might include conditions such as a history of smoking, obesity, or congestive heart failure.



Fetal Death Data

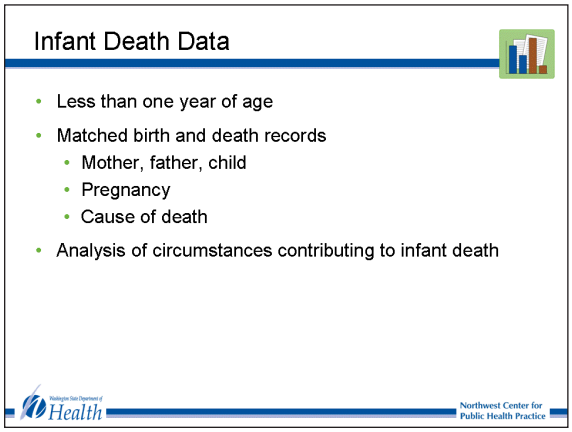
Fetal death data include fetal deaths occurring in Washington State and fetal deaths to Washington State residents occurring out-of-state. Reporting is required when the gestational age of the fetus is 20 weeks or more. These records include information about the mother, father, course of the pregnancy, and the cause of the fetal death.

Fetal death records also provide a legal record of death and information about prenatal care.

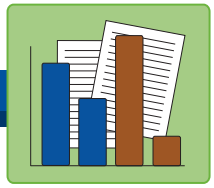


Infant Death Data

Infant mortality data are a special subset of death records because each death certificate for a child less than one year of age is matched to the data of that child's birth certificate. These matched records provide information about the mother, father, child, pregnancy, and cause of death. These data allow for the analysis of circumstances that may have contributed to infant deaths.

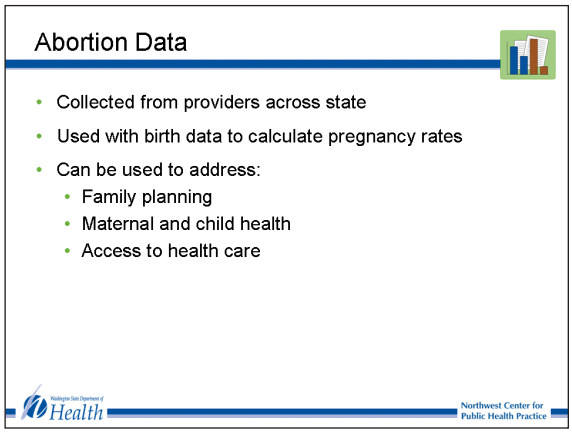


Transcript



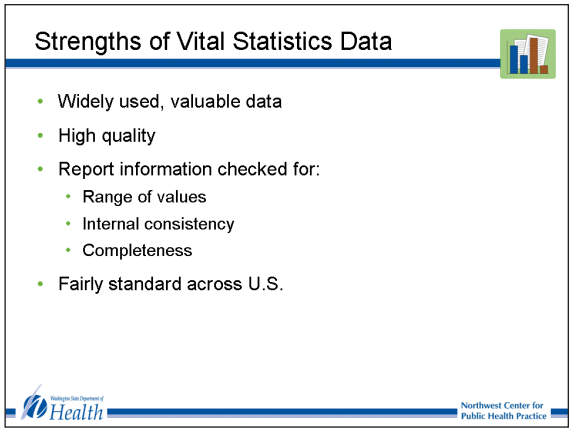
Abortion Data

Abortion data are collected from health care providers across the state. When combined with birth data, they can be used to calculate pregnancy rates. Data on pregnancy and abortion rates can be used to address family planning, maternal and child health, and other concerns about access to health care.



Strengths of Vital Statistics Data

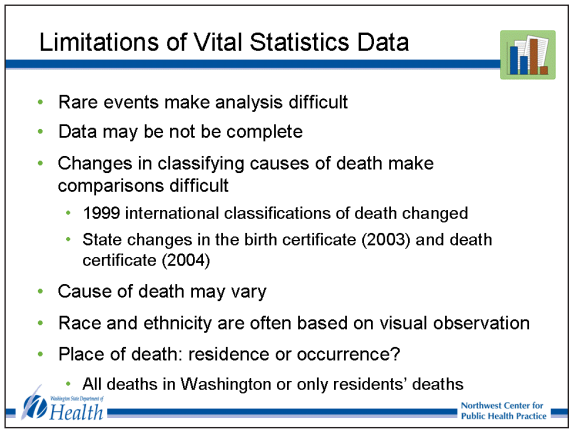
Vital statistics records provide some of the most widely used and valuable data for public health assessment. They represent complete records on births, deaths, and infant deaths for Washington State residents. The data are of high quality because, when possible, reported information is checked to see if it is within a reasonable range, if there is internal consistency among items, and if the information is complete. In addition, vital records data are fairly standard across the U.S., allowing for cross-state comparisons.



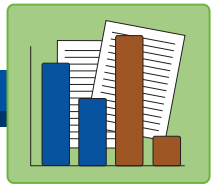
Limitations of Vital Statistics Data

Vital records data have some limitations, of course. Some of the events captured in the data, such as infant deaths, are relatively rare, which make analysis of the data difficult because the small numbers result in unreliable rates. (Modules 2 and 3 examine the problems of small numbers and how to deal with them.) Also, fetal death data may not be complete so it is important to remember that the data may underestimate the number of fetal deaths.

Although official standards maintain the consistency of the data for the most part, changes to those standards do occur. These changes make comparisons with earlier years of data difficult. For example, in 1999, the international system for classifying cause of death was updated, which changed the definitions of some causes of death.



Transcript



More recently, changes to the way information is collected on birth and death certificates were implemented in Washington State.

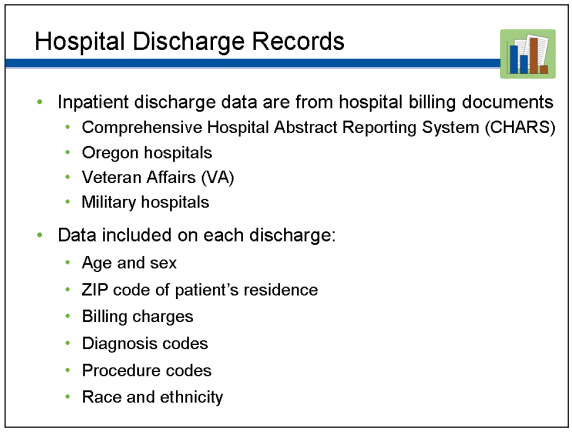
Also, even within the same classification system, the determination of cause of death may vary depending on the practitioner filling out the death certificate. Cause of death is a legal as well as medical opinion and is subject to the best information available to the physician, medical examiner, or coroner certifying the cause of death. In addition race and ethnicity are often determined by funeral directors, based on visual observation, which can lead to misclassification.

One final note of caution for all vital records data has to do with the analysis of "place." All of these events may occur in places other than an individual's residence. However, they are all generally analyzed by residence. Always pay attention to whether the data include all deaths that occurred in Washington State, or only deaths of Washington State residents.

Let's pause now so you can check your understanding of the material we've just covered.

Practice: Vital Statistics Data

Hospital Discharge Records



Hospital discharge record data are available from the inpatient discharge database in Washington State, which is built from hospital billing documents. The data are collected by the Center for Health Statistics at Washington State Department of Health.

Washington's hospital discharge data system includes data from acute care hospitals in Washington State. This database is primarily from the Comprehensive Hospital Abstract Reporting System (or CHARS), but additional hospitalization data are also available for public health assessment purposes from Oregon hospitals, the Veterans Affairs (or VA), and military hospitals. The data collected on each hospital discharge include:

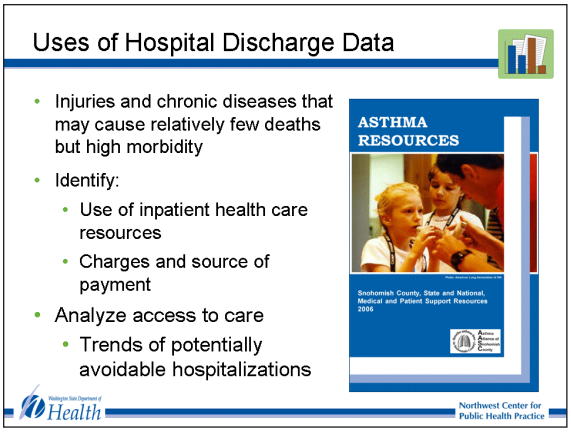
Age and sex of the patient, ZIP code of the patient's residence, billed charges, diagnosis codes, procedure codes, and, since 2007, race and ethnicity.

Often in assessment we want to know about the trends and rates of conditions that are costly or severe and that might be prevented with earlier prevention strategies. Asthma is a good example of a condition that should not end up in a hospi-

Transcript

talization. Since the goal of many asthma coalitions or programs in local health departments and states is to prevent severe asthma, the CHARS data can be used to identify sub-groups or areas that have high hospitalization rates to target intervention strategies.

Uses of Hospital Discharge Data

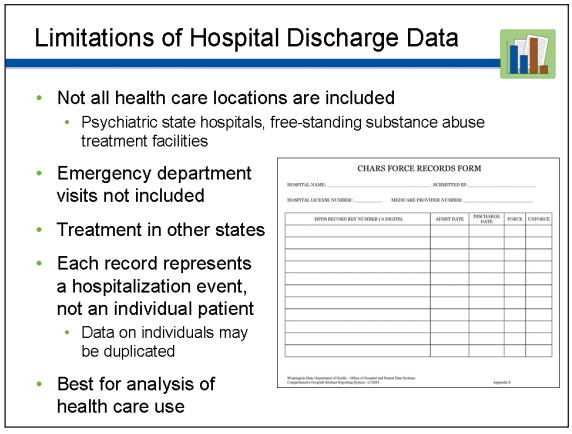


Hospital discharge data is the primary source of data on injuries and chronic diseases that may cause relatively few deaths but high morbidity, or illness. In public health we use the data to identify the use of inpatient health care resources, charges for hospital care, and source of payment.

We also use the data to analyze access to health care by examining trends of potentially avoidable hospitalizations.

In an asthma resource guide we developed in Snohomish County, we compared the hospitalization rates of asthma in children to the rates in adults. The rate was twice as high in children, which emphasized the need to diagnose and provide early prevention strategies for children to avoid hospitalizations.

Limitations of Hospital Discharge Data



As you might guess, an important limitation of hospital discharge data is incompleteness.

For one thing, not all health care locations are tracked. For example, psychiatric state hospitals are not included. Neither are free-standing substance abuse treatment facilities. In addition, visits to emergency departments that don't result in a hospital admission aren't included. And finally, Washington residents treated in hospitals in states other than Oregon are not included.

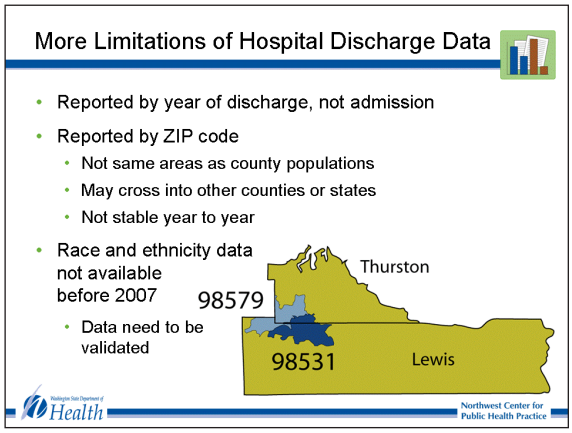
Another problem is that each record in the hospital discharge data represents a hospitalization event, not an individual patient. An individual, for example, may have multiple admissions in the same year for the same cause, such as asthma. Rates of hospitalization for a particular cause that are calcu-

Transcript

lated from these data, therefore, don't represent the number of persons hospitalized. This means data on individuals may be duplicated. However, it is possible to request unduplicated data.

Although hospital discharge data are frequently used as a substitute measure for estimating prevalence, they are best used for the analysis of health care use.

More Limitations of Hospital Discharge Data



We face three other problems when working with hospital discharge data.

Hospitalization events are reported by the year of discharge—not the year of admission.

Hospital discharge data are reported by the ZIP code of residence. In order to calculate county rates, ZIP codes are aggregated to approximate county boundaries. However, ZIP code-based county aggregations do not always equal county populations because some ZIP codes cross into other counties or states. In addition, ZIP code boundaries aren't stable from year to year.

And finally, the hospital discharge data system only started collecting race and ethnicity in 2007, so historical data aren't available for these categories, and they also need to be validated.

Asthma offers an example of an important limitation of hospital data. Using just hospitalization data would not give us a complete picture of the number of asthma cases that have severe or are potentially preventable episodes. Many times asthmatics go to the emergency room, but are not hospitalized. As a result, we underestimate the number of asthma cases that need earlier prevention strategies.

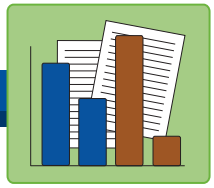
Let's pause now so you can check your understanding of the material we've just covered.

Practice: Hospital Discharge Data

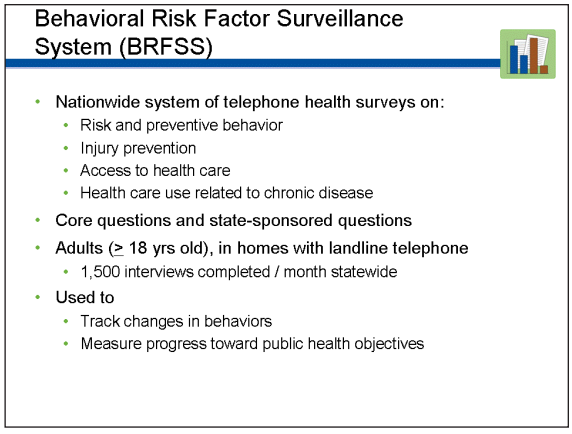
Behavioral Risk Factor Surveillance System (BRFSS)

The Behavioral Risk Factor Surveillance System (or BRFSS) is a nationwide, state-based system of telephone health surveys with adults. It gathers information on:

Transcript



- Risk and preventive behavior, such as diet, exercise, and current chronic conditions
- Injury prevention, such as seatbelt use
- Access to health care, such as having health care insurance, and
- Health care use related to chronic disease, such as cancer screenings



The survey is conducted in 50 states, the District of Columbia, and three territories.

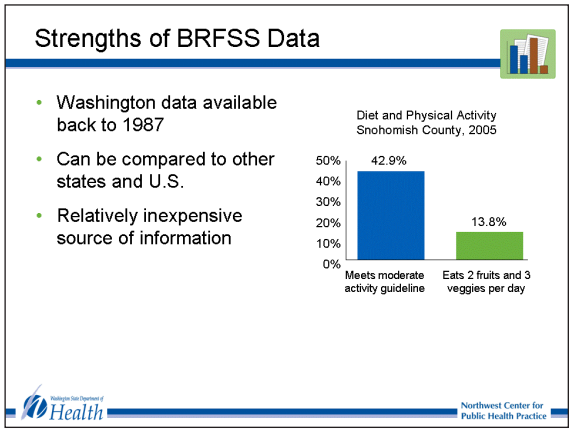
The survey questionnaire includes a core set of questions used by all states and an additional set sponsored by each state, according to its needs.

Participants are selected by randomly sampling households with a landline telephone. Once selected, an adult aged 18 or older living in the household is chosen to respond to questions.

In Washington State, approximately 1,500 interviews are completed each month.

Data from BRFSS are used to track changes in health-related behaviors and measure progress toward public health objectives in reducing chronic disease prevalence. For example in Snohomish Health District, because of the increasing trends in obesity, we identified it as a priority to address in our county.

Strengths of BRFSS Data



BRFSS data for Washington State are available back to 1987. This enables public health practitioners and researchers to document trends, measure progress, and identify emerging health issues. You can find a link to the Washington State BRFSS Web site in the Resources at the end of this module.

Because of the core question set used nationally, the results of many questions can be compared to other states' and to the total U.S. population.

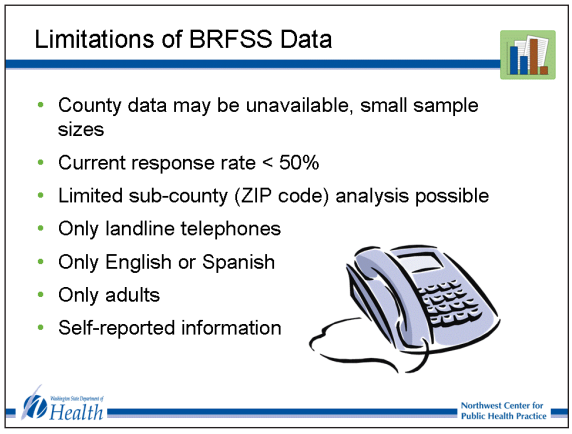
BRFSS is a relatively inexpensive way for states to gather a wide range of information about the population's health.

Having information about chronic diseases, prevention and risk behaviors, and demographic information provides health department data to help target intervention strategies. In addition to the increasing trends in obesity in Snohomish County,

Transcript

we observed low levels of physical activity and fruit and vegetable consumption. This information, and other information in BRFSS, is a great resource to aid local and state health departments to plan and to identify priorities and target populations or communities for intervention strategies.

Limitations of BRFSS Data



BRFSS data, however, have many of the limitations of any population survey. Even though the state began gathering BRFSS data in 1987, county data may not be available, and sample sizes in earlier years were smaller. Over time, questions and methods have also changed.

It is important to take these limitations into consideration when analyzing your data.

Since the current response rate is less than 50%, you need to be aware of potential biases in the data. Also, because of the small sample size, only limited sub-county analysis on the ZIP code level is possible. In addition, counties with smaller populations will have a smaller number of cases. Although a few years of data can be combined to provide more reliable data at the county level (as we discussed in Module 3), sub-county analysis in smaller counties will be limited.

Since only those who have a landline telephone are called, those with cell phones as their only phone are not represented in the data. Young adults, who are more frequent owners of only cell phones, are the most likely to be underrepresented. Also, households without any telephone service are not included in BRFSS. The state and CDC are now doing research to evaluate possible survey bias related to cell phones.

In addition, the survey covers only adults who speak English or Spanish. In our increasingly diverse population, this may be a limitation, especially for counties that have more minorities.

And finally, all information is from adults and is self-reported by the respondents.

Practice: BRFSS Data

Healthy Youth Survey (HYS)

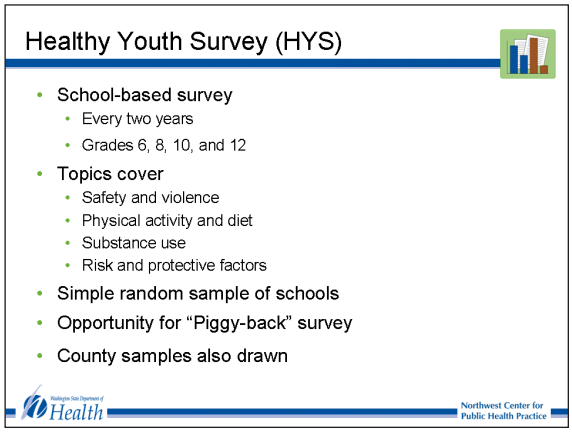
The Healthy Youth Survey is a school-based health and risk behavior survey that provides information about adolescents in Washington State. Every other year, in

Transcript

even years, students in grades 6, 8, 10, and 12 of Washington public schools are surveyed. The survey has been conducted every other fall since 2002.

Students answer survey questions about safety and violence, physical activity and diet, alcohol, tobacco, and other drug use, and related risk and protective factors.

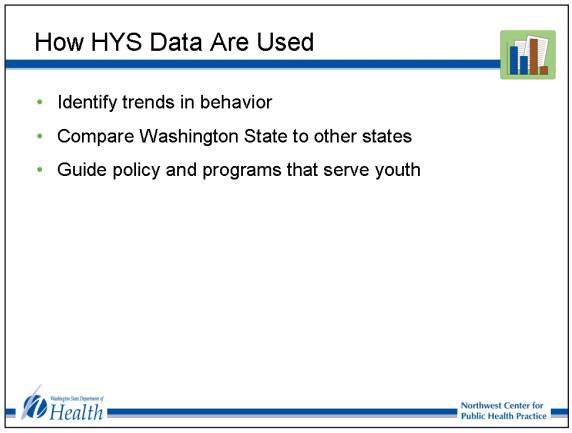
Respondents are selected through a simple random sample of public schools. Schools that aren't selected to participate are offered an opportunity to do a "piggy-back" survey, or in other words, to be included in the survey, to obtain local data. In large counties, such as King, Spokane, and Snohomish, additional county samples are drawn.



How HYS Data Are Used

Data from the Healthy Youth Survey can be used to identify trends in the behavior of Washington adolescents over time.

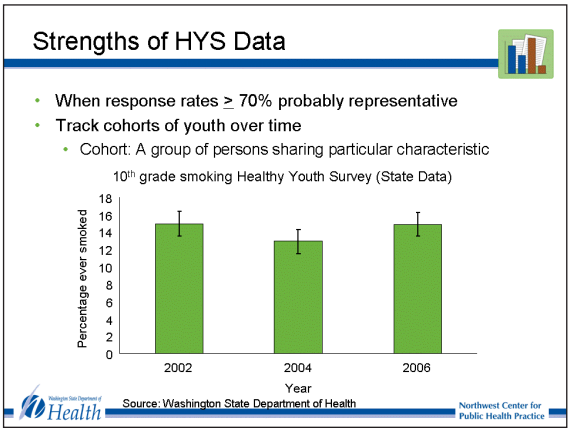
The state-level data can be used to compare Washington to other states that do similar surveys as well as well as to the nation. Public health, schools, and others use this information to guide policy and programs that serve youth. In Snohomish Health District, for example, we used HYS data to educate Healthy Community Advisory Committees and assess the needs of the community to set goals and priorities.



Strengths of HYS Data

Response rates for HYS are calculated based on enrollment in the schools asked to participate. When response rates are 70% or greater, the results are probably representative of Washington students in that grade group.

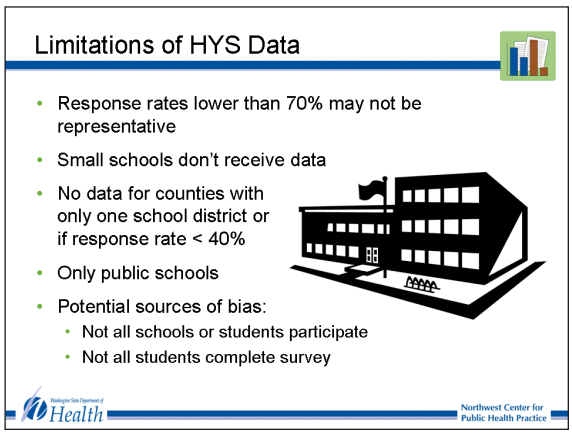
Data collection from these grades every two years enables communities and state agencies to



Transcript

watch "cohorts" of youth over time. A cohort, by the way, is a group of persons sharing a particular statistical or demographic characteristic. So the students in grade 6 in 2006 are considered to be the same cohort in grade 8 in 2008, even though the individual students may not be the same.

Limitations of HYS Data



As with the other data sources we've discussed, the Healthy Youth data have a number of limitations.

For one, response rates lower than 70% may not be representative of all students in that grade level in the state.

Also, although all schools are encouraged to participate and contribute to the overall district or county-level estimates, small schools with fewer than 15 students in a grade don't receive their results, in order to protect the confidentiality of the students.

Results are not reported for counties with only one school district, unless the school district agrees. Also results are not reported for specific grades within a county if the response rate was less than 40 percent. This is primarily because of the lack of reliability of the data with such a low response rate. In addition, it is important to be aware that data for a school or a district cannot be presented without their approval.

Another limitation is that the survey is conducted only in public schools. So, children in private schools or being home schooled as well as youth who have dropped out of school are not represented in the data.

Finally, as with all surveys, several other potential sources of bias exist in the data. For example, not all schools or students chosen to be part of the sample participate, and, not all students complete the survey.

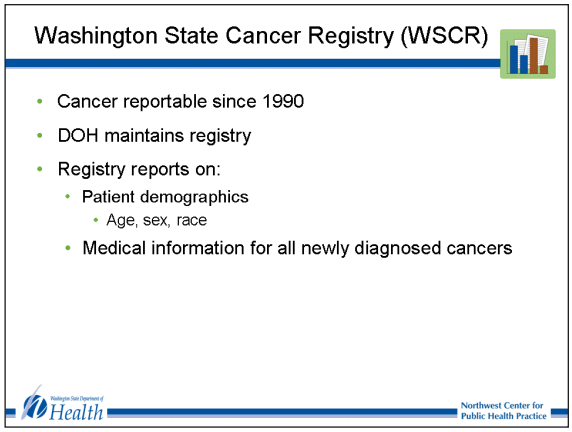
Let's pause now so you can check your understanding of the material we've just covered.

Practice: HYS Data

Transcript

Washington State Cancer Registry (WSCR)

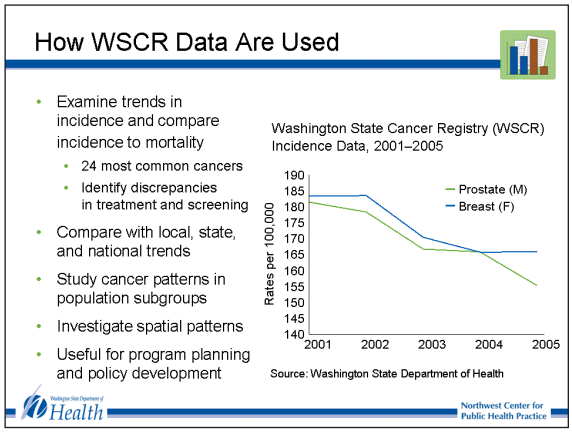
I talked earlier about reportable diseases registries, and the cancer registry is another example of a reportable condition registry. Cancer has been a reportable condition in Washington since 1990. The Department of Health maintains the cancer registry. The registry reports data on patient demographics, such as age, sex, and race, and on medical information for all newly diagnosed cancers.



How WSCR Data Are Used

Washington State Cancer Registry data are available for examining trends in incidence and comparing incidence to mortality associated with the 24 cancers most frequently reported in the state. This comparison is useful for assessing discrepancies in treatment and screening practices.

You can find a link to the Washington State Cancer Registry Web site in the resources at the end of this module.

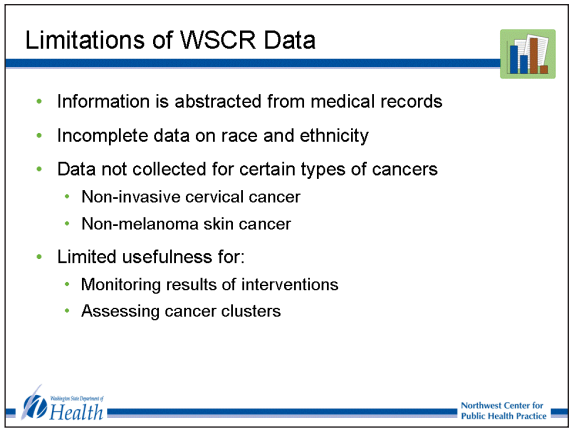


Because comparable data exist in many states, Washington State data can be compared to other states and the nation. Cancer patterns in population subgroups within the state can be also studied, as well as spatial patterns and other similar characteristics. Combined with information on cancer prevention, early detection, and treatment, these data are useful for program planning and policy development aimed at reducing the burden of cancer.

In Snohomish Health District for example, when we identified the top causes of mortality, cancer increased to the point where it is now the most common cause of death, and it is approximately equal to heart disease. To identify where specific prevention strategies might be important, we used the cancer registry to determine the number of new cases of cancer to identify the most common kinds of cancer. Breast cancer in females and prostate cancer in males have the highest incidence. Combined with the mortality information, this sends an important message to providers that screening and early diagnosis are important prevention methods to continue.

Transcript

Limitations of WSCR Data



Data from the cancer registry have a number of limitations.

The information in the Washington State Cancer Registry data set is abstracted from medical records and depends on the reporting capabilities of the individual facilities making the reports. This may lead to incomplete data.

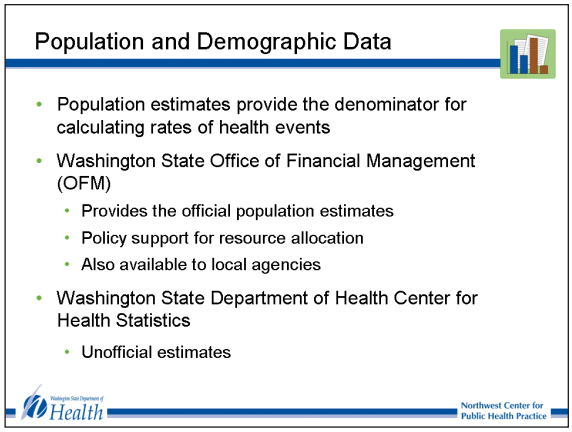
In addition, as with many data sets, race and ethnicity are self-reported, and these data are frequently incomplete. As a result, specific minorities may be under reported. Also, data are not collected for certain types of cancers such as non-invasive cervical cancer and non-melanoma skin cancer.

Finally, because of the length of time for cancer to develop and be diagnosed, the data are of limited usefulness in monitoring the effect of interventions aimed at primary prevention. This delay —along with the small number of cases in a given community—also limits the ability to assess apparent clusters of cancer in communities.

Let's pause now so you can check your understanding of the material we've just covered.

Practice: WSCR Data

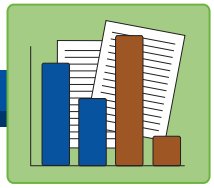
Population and Demographic Data



Population estimates and demographics are crucial for public health assessment because they provide the denominator figures necessary for calculating rates of health events. If you'll recall, in order to calculate a rate, such as prevalence or incidence, we need to know the total population of interest. (Modules 2 and 3 in this series discuss various important rates.)

The Washington State Office of Financial Management, or OFM, provides several sources of

Transcript



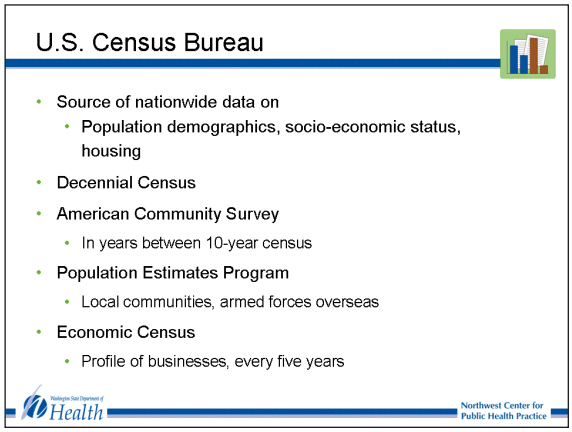
data on the state population and its demographics, as well as the official estimates of the population of counties and cities in Washington State.

The purpose of these estimates is to provide policy support and information used in resource allocation to state agencies, including the state Department of Health, but local departments can also use the data.

The Center for Health Statistics at the Washington State Department of Health produces "unofficial" population estimates at the state and local level for assessment purposes.

U.S. Census Bureau

The most important source of data on population demographics, socio-economic status, and housing, of course, is the U.S. Census Bureau. In addition to the more well-known decennial census, the Census Bureau also provides access on its Web site to the data from several other surveys, including the American Community Survey, the Population Estimates Program, and the Economic Census.



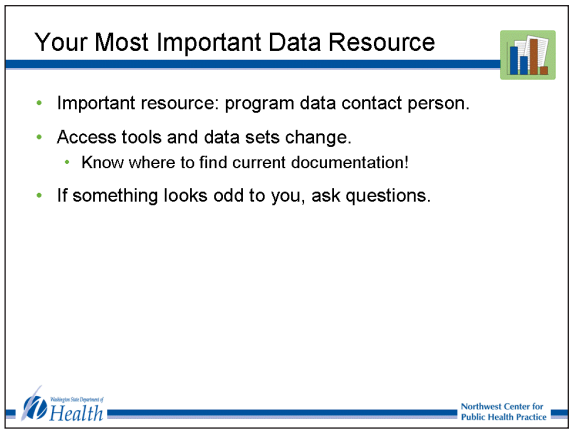
The American Community Survey samples a small percentage of the American population on a rotating basis, to help fill in the gaps between the 10-year census.

The Population Estimates Program provides estimates not only for local communities by demographic variables (such as, age, sex, race, and Hispanic origin), but also includes estimates for the armed forces overseas.

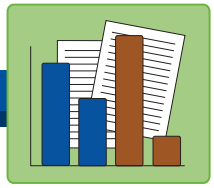
The Economic Census provides a profile of American businesses every five years, at both the national and the local level.

Your Most Important Data Resource

When using these data sources or any others, one of the most important resources is the data provider's program data contact person. Data access tools change, and datasets themselves change over time, so you should also know where to find the current documentation on a data source.



Transcript



Data users sometimes discover inconsistencies in the data. If something looks odd, don't hesitate to ask questions of the data provider's contact person.

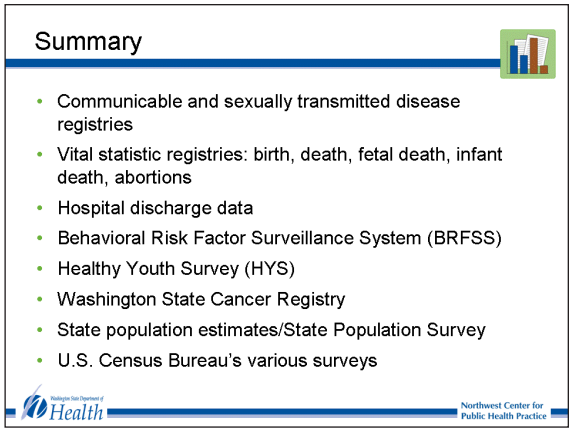
Let's pause now so you can check your understanding of the material we've just covered.

Practice: Finding the Data You Need

Summary

To summarize, in this module we discussed a variety of data sources:

- Communicable Disease and Sexually Transmitted Disease registries,
- Vital statistic registries for birth, death, fetal death, infant death and abortions,
- Hospital discharge data, which you can find in the Comprehensive Hospital Abstract Reporting System, or CHARS,
- The Behavioral Risk Factor Surveillance System, or BRFSS,
- The Healthy Youth Survey, or HYS,
- The Washington State Cancer Registry,
- State population estimates and the State Population Survey, and finally,
- The U.S. Census Bureau's various surveys.



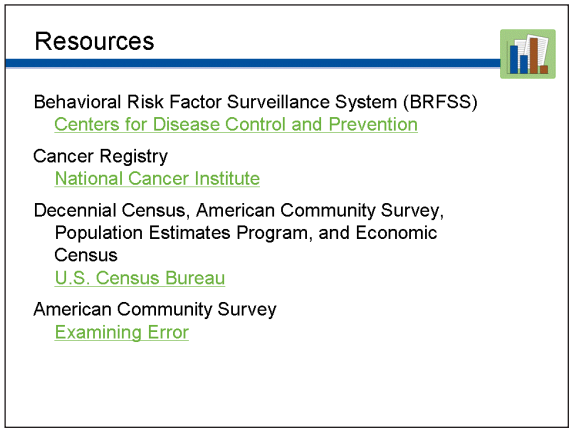
Even though these data sets have a variety of limitations, they are very valuable and can be used to set priority areas, identify prevention activities, identify target populations, and educate communities about their health.

Resources

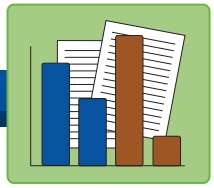
If you would like to learn more about the concepts in this module, you might want to explore some of the resources listed here.

Behavioral Risk Factor Surveillance System (BRFSS)

Centers for Disease Control and Prevention www.cdc.gov/brfss



Transcript



Cancer Registry
Centers for Disease Control and Prevention, National Cancer Institute
www.cancer.gov

Decennial Census, American Community Survey, Population Estimates Program, and Economic Census

U.S. Census Bureau www.census.gov

American Community Survey

Examining Error www.esri.com/news/arcuser/0708/demoarticle.html

Now, if you're ready, please go on to the final test.